

Truthful Mechanisms for Multi Agent Self-Interested Correspondence Selection

Nan Zhi¹, Terry R. Payne¹, Piotr Krysta¹, and Minming Li²

¹ The University of Liverpool, Liverpool, United Kingdom
{N.Zhi, T.R.Payne, P.Krysta}@liverpool.ac.uk

² City University of Hong Kong, China
minming.li@cityu.edu.hk

Abstract. In the distributed ontology alignment construction problem, two agents agree upon a meaningful subset of correspondences that map between their respective ontologies. However, an agent may be tempted to manipulate the negotiation in favour of a preferred alignment by misrepresenting the weight or confidence of the exchanged correspondences. Therefore such an agreement can only be meaningful if the agents can be incentivised to be honest when revealing information. We examine this problem and model it as a novel mechanism design problem on an edge-weighted bipartite graph, where each side of the graph represents each agent's private entities, and where each agent maintains a private set of valuations associated with its candidate correspondences. The objective is to find a matching (i.e. injective or one-to-one correspondences) that maximises the agents' social welfare. We study implementations in dominant strategies, and show that they should be solved optimally if truthful mechanisms are required. A decentralised version of the greedy allocation algorithm is then studied with a first-price payment rule, proving tight bounds on the Price of Anarchy and Stability.

Keywords: Decentralised Ontology Alignment · Multi-Agent Systems

1 Introduction

Within open, distributed environments, agents may differ in the way they model a domain, and may assume different logical theories or *ontologies* [13]. This can result in the existence of numerous models that, despite modelling a similar domain, are themselves semantically heterogeneous, and thus not interoperable. These ontological models can be reconciled by computing an *alignment*: i.e. a set of *correspondences* (*mappings*) stating logical relationships between the entities in the different ontologies [10]. Two agents may be able to communicate and thus transact if their individual ontologies cover the same domains, and if a meaningful alignment can be found.

Various static (single-shot) and dynamic approaches [15, 19] have explored how agents can propose, and exchange candidate correspondences with the goal of aligning their respective ontologies. In many cases, agents acquire knowledge

of different candidate correspondences from a variety of sources, or through negotiation with other agents. These candidate correspondences may have an associated *weight*, which may reflect the utility, significance, or simply the confidence that an agent has in the correspondence. Furthermore, in adversarial scenarios, the agents may not wish to disclose their private weights, and may lie when stating their preferences.

As the composition of different subsets of correspondences can result in different alignments, the challenge in negotiating a mutually acceptable alignment is that of selecting and proposing correspondences that result in a preferred alignment that satisfies the aims of both agents. Furthermore, some correspondences may map a single entity in one ontology to different entities in other ontologies (which can compromise the integrity of the resulting logical model), and therefore the outcome should ideally be injective (i.e. a matching).

In this paper, we take a mechanism design based approach to investigate and analyse theoretically the problem from a centralised perspective (*Dominant Strategies*), where the problem is characterised as a social welfare maximising matching setting with an additive valuation function. To model this from a mechanism design perspective, we use the term “*payment*” to refer to the agent’s view of the correspondence’s weight. We show that for a deterministic mechanism with payment, the only truthful mechanism is *maximal-in-range* (defined within Section 4), and any truthful mechanism which is not optimal can do no better than an approximation ratio of 2. Given our results on truthful centralised mechanisms, either the problem should be solved optimally (though costly) or strong lower bounds should be found for the approximation ratios of truthful mechanisms. We have also explored an implementation in Nash Equilibria [25] to efficiently approximate mechanisms for matching using the greedy allocation mechanism.

In Section 2, the challenges of selecting correspondences for injective alignments are discussed from a centralised and decentralised standpoint. In Section 3, the Ontology Alignment Selection problem is formalised, and examined from a decentralised (two agent) perspective. The problem is then analysed as a Mechanism Design game with payment in Section 4. A Greedy Algorithm is studied as a means of finding an approximate Nash Equilibria solution, and its properties are formally proved (Section 5). This is followed by a discussion and related work in Section 6, before concluding in Section 7.

2 Background

To date, the ontology alignment community has proposed many diverse approaches that *align* ontologies in order to find sets of correspondences between the ontology pairs.³ However, most approaches rely on the ontologies being fully shared with some alignment algorithm [10, 27] which attempts to find correspondences between entities. Alignment approaches usually initiate the process of

³ For a comprehensive overview of the different approaches, we refer the reader to the *Ontology Alignment Evaluation Initiative* - <http://oaei.ontologymatching.org>

identifying correspondences (mappings) by computing a *similarity matrix* (lexical, structural or a combination of these) between all the entities in the two ontologies that are being aligned [10, 22]. This produces a number of different mappings involving the same entities from which an *injective* (one-to-one) alignment needs to be extracted (i.e. correspondences for which to each entity from the source ontology corresponds only one entity in the target ontology).

Typically, most alignment approaches model the alignment as a bipartite graph, and thus select an injective alignment by finding a *matching* or independent edge set in the graph, such that the set of edges (i.e. correspondences) have no common vertices (i.e. no entity in one ontology is mapped to more than one entity in the other ontology, and vice versa). This assumes that each edge (or correspondence) is weighted such that the weight represents the *quality* or *desirability* of the correspondence. The two most common methods used to select a *matching* are: 1) to find a global optimal solution (which is equivalent to the *Assignment Problem*) using algorithms such as the Hungarian method [18];

or to find a sub-optimal, but *stable* solution using algorithms such as Gale & Shapley’s Stable Marriage algorithm [14]. Solutions to the assignment problem identify correspondences that maximise the sum of the weights (i.e. they assume some *objective function* that maximises *social welfare*), as opposed to the similarity of each pair of entities. This is illustrated in Figure 1, where two correspondences are selected by maximising the weights; in this case the weights associated to the two correspondences $\{e_1, e_3\}$ are $1 + 1 = 2$. As ontologies can vary greatly in size, with several in the Bio-Medical domain possessing tens of thousands of entities [16], techniques such as the Hungarian method can become computationally costly ($O(n^3)$ for its most efficient implementation). Thus, sub-optimal approximate algorithms such as a greedy matching algorithm [22] or a variant from the family of Stable Marriage algorithms [14] are used that select a sub-optimal set of correspondences in those cases when a *stable* solution is sufficient. This can result in a different alignment that emphasises the weights of individual correspondences; given the example in Figure 1, a greedy algorithm would generate an alignment with a single correspondence, e_2 , as its weight is greater than either e_1 or e_3 , resulting in a sub-optimal total weight of $1 + \epsilon$.

A similar problem arises in decentralised settings, where agents negotiate over a set of (partially observable) correspondences to agree upon a mutually acceptable alignment [3, 6, 9, 15, 19, 26], often based on the aims or goals of the agents that may own or utilise them. As no single alignment approach can provide a panacea for all ontology pairs, agents are left with the problem of either: 1) *selecting* a suitable alignment approach from the plethora that exist; or 2) *assembling* alignments from a subset of relevant, candidate correspondences; for example using an ensemble approach. This latter case occurs if agents have

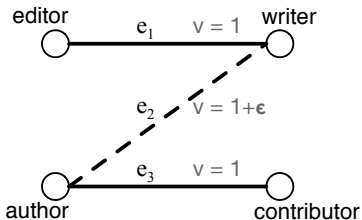


Fig. 1. Centralised example with two solutions: $\{e_1, e_3\}$ and $\{e_2\}$.

access to correspondences from shared repositories [19] or garnered from previous transactions with other agents. Furthermore, alignments with different constituent correspondences may be semantically equivalent with respect to one of the agent’s ontologies and aims (due to the logical theory underlying each agent’s ontology) but may have a different meaning to another.⁴ As the agent may have preferences over the choice of correspondences used (e.g. due to privacy concerns [12, 23]), agents can have a preference order over the resulting alignments within the same equivalence class. Hence, for *self-interested* agents, this task becomes one of selecting a mutually acceptable subset of preferred ontological correspondences.

The resulting alignment will typically be dependent on the *value* that each agent associates to each correspondence. Whilst this is uncontroversial in centralised systems, approaches that are decentralised (i.e. where agents may differ in the value they ascribe to a correspondence) are subject to strategic manipulation; i.e. agents may lie about the true value of a correspondence to ensure that the final alignment includes their preferred correspondences. The value that each agent assigns to each correspondence (i.e. its *private valuation*) relates to how useful this edge is in resolving a query or achieving a task, and in turn, the potential payoff the agent can obtain from performing a task. Note that this is not the same as the confidence the agent has in the edge (based, for example from some form of linguistic similarity metric over the concept labels). For example, an agent may know of two correspondences in the publishing domain `{writer, editor}` and `{writer, author}`. Both are viable correspondences, depending on the task (e.g. for a conference proceedings and monograph respectively), but an agent may assign different valuations to each correspondence based on some preference; for example the agent can increase its payoff by resolving queries or performing tasks (by providing a service to its peers) pertaining to monographs. Conversely, it may have a low valuation for other correspondences for which it has little preference (e.g. `{writer, publisher}`). However, within a service landscape where several agents (providing services) may compete to perform a task for a requesting agent, they may not wish to disclose the true value of this payoff. This can potentially lead to agents strategically manipulating the combined value of sets of correspondences, in order to maximise their individual payoffs; potentially resulting in semantically compromised correspondences being selected, which may then prevent the query or task from successfully completing. Thus, in an ideal setting, the agents should be incentivised to adopt strategies that result in alignments that benefit both agents; i.e. find solutions that lie within a *Nash Equilibrium* [25].

3 The Decentralised Alignment Construction Problem

We consider the Alignment Construction Problem given the following setting in which there are two agents $i \in \{L, R\}$ (the *left* agent and *right* agent), where each

⁴ A classic example of terminological difference exists with the term “*football*”, which has a different meaning depending on whether the reader is from the US or the UK.

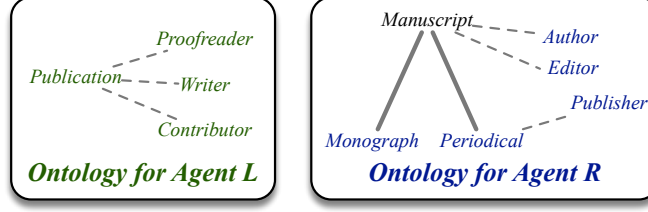


Fig. 2. Ontology fragments \mathcal{O}_L (left) and \mathcal{O}_R (right). The solid line denotes the *isa* class hierarchy relation, whereas the dashed line indicates property relations between classes (note that property names are not given).

agent i possesses a private ontology \mathcal{O}_i that includes the *named concepts* (i.e. *entities*)⁵ $\mathbf{N}_i^C \in \mathcal{O}_i$ to be aligned. The alignment is modelled as an edge-weighted bipartite graph $G = (U \cup V, E)$, where the vertices of U and V correspond to the entities in the agents' individual ontologies $U = \mathbf{N}_L^C$ and $V = \mathbf{N}_R^C$ respectively, and the edges $e \in E$ correspond to the candidate correspondences. A matching M is a subset of E such that $e \cap e' = \emptyset$ for all $e, e' \in M$ with $e \neq e'$; i.e. no two edges have a common vertex. Each agent $i \in \{L, R\}$ has a non-negative valuation function for different matchings M , denoted $v_i(M)$, where $v_i : M(G) \rightarrow \mathbb{R}^+$, which is additive; i.e. $v(S) + v(T) = v(S \cup T)$ such that $S \cap T = \emptyset$ for all $S, T \subseteq M$, and $M(G)$ is the set of all matchings in a graph G . Each agent i also has a valuation function $v_i : E \rightarrow \mathbb{R}^+$ to represent the value $v_i(e)$ it privately ascribes to the edge e . The combined value for an edge e is therefore given as $v(e) = v_L(e) + v_R(e)$. Note that $v_i(M) = \sum_{e \in M} v_i(e)$ for every agent $i \in \{L, R\}$, and $v(M) = \sum_{i \in \{L, R\}} v_i(M)$ is the combined value for the matching M .

The goal is to establish an alignment which is equivalent to a matching M that maximises $\sum_{e \in M} v(e)$; i.e. find a set of edges whose sum of weight is maximal. This problem, known as the *Assignment Problem*, is typically solved optimally using Kuhn's *Hungarian Algorithm*[18]. In a distributed negotiation setting, the valuation function v_i can be regarded as the agents' true valuation, or *type* that it attributes to each matching. Furthermore, we use v to represent the combined type profile for both agents, such that $v = \{v_L, v_R\}$, where v_i is the type profile for agent i , and similarly, b denotes the combined bid profile for both agents (see Section 3.1 below for details on bids), such that $b = \{b_L, b_R\}$, where b_i is the bid profile for agent i . We will also introduce the following useful notation: $b_i^e = b_i(e)$ and $v_i^e = v_i(e)$ for any $i \in \{L, R\}$ and $e \in E$.

Consider the Bookseller scenario illustrated in Figure 2 for agents L and R , where each agent possesses a simple ontology fragment within the *publishing* domain. Agent L models the class entity *Publication* in \mathcal{O}_L with three property relations (unnamed in this example) to three other entities: *Proofreader*, *Writer* and *Contributor*. The other agent models the same domain but with entities from

⁵ We follow the standard practice of restricting ourselves to correspondences between named concepts within the respective ontologies [10], and omit the discussion of the property relations between entities within each ontology.

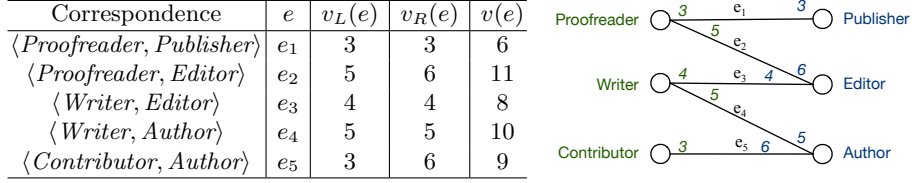


Fig. 3. The individual weights for different correspondences (left) that map entities from \mathcal{O}_L to those in \mathcal{O}_R . The combined edges $v(e)$ appear in the final column. The resulting graph (right) has two possible matchings: an optimal matching $M_{opt} = \{e_1, e_3, e_5\}$ where $v(M_{opt}) = 23$, and a stable matching $M_{stable} = \{e_2, e_4\}$, where $v(M_{stable}) = 21$.

\mathcal{O}_R . The class *Manuscript* has two subclasses in particular: *Monographs* (i.e. a specialist work by a single or small number of authors) and *Periodicals* which are edited volumes containing numerous articles (written by different authors). Both subclasses have properties to the concepts *Author* and *Editor* (inherited from *Manuscript*), whereas *Periodical* also has a property to the concept *Publisher*. The table in Figure 3 (left) lists candidate correspondences between entities in L 's ontology, and those in the ontologies of agents R , complete with each agents private valuation function for each correspondence e , and a label e_i .

3.1 Alignment Construction with Payment

To model this problem from a mechanism design perspective (or a game, where two agents cooperate with each other to find a resulting alignment), we consider the notion of agents declaring a value for each correspondence. As this value could differ from their private value (because each agent may be behaving strategically to manipulate the outcome), we refer to the declarations as *bids*. For this reason, we consider a mechanism with *payments*. We define a direct revelation mechanism $\mathcal{M}(\mathcal{A}, \mathcal{P})$, which is composed of an allocation rule \mathcal{A} to determine the outcome of the mechanism (i.e. this determines which of the correspondences are selected for the resulting alignment), and a payment scheme \mathcal{P} which assigns a vector of payments to each declared valuation profile. The mechanism proceeds by eliciting a bid profile b_i from each agent i , and then applies the allocation and payment rules to the combined bid profiles to obtain an outcome and payment for each agent. As an agent may not want to reveal its type (i.e. its true value), we assume that b does not need to be equal to v .

The utility $u_i(b)$ for agent i given a bid profile $b = (b_L, b_R)$ and mechanism \mathcal{M} is based on the allocation rule \mathcal{A} and the payment scheme \mathcal{P} over the outcome of $\mathcal{A}(v)$ (i.e. a matching or allocated set M), and can be written as $u_i(\mathcal{A}(b)) = v_i(\mathcal{A}(b)) - \mathcal{P}_i(b)$. For an implementation in Nash Equilibria (see Section 5), we assume a *first-price* payment rule, such that an agent is charged its declared bid $b_i(M)$ for any allocated set M . Our objective function maximises the social welfare SW given both agents' bids (generating either optimal or approximately optimal solutions); i.e. $SW(\mathcal{A}(b), v) = \sum_{e \in \mathcal{A}(b)} v(e)$. A (deterministic) mechanism \mathcal{M} is called *truthful in dominant strategies* or *incentive*

compatible if, for any agent $i \in \{L, R\}$, we have $u_i(\mathcal{A}(v_i, b_{-i})) \geq u_i(\mathcal{A}(b_i, b_{-i}))$ for any bid profile b_i of agent i and any bid profiles b_{-i} of the other agents.⁶

3.2 Nash Equilibria

Different types of Nash Equilibria may exist, depending on the strategy adopted by the agents. The bid profile b forms a *Pure Nash equilibrium* if, for both agents, there exists no other bid profile b'_i achieving a higher utility, i.e., $\forall b'_i, u_i(b_i, b_{-i}) \geq u_i(b'_i, b_{-i})$. As no agent can obtain a higher utility by deviating from b ; they can do no better than to select alignments that result in a Nash Equilibrium [25].

We also permit a randomised strategy function which can result in a *Mixed Nash equilibrium*. Given the probability distribution $\omega_1, \dots, \omega_n$ over the declarations, and any function f over the space of declaration profiles, we can state $\mathbb{E}_{b \sim \omega}[f(b)]$ for the expected value of f over declarations chosen according to the product distribution $\omega = \omega_1 \times \dots \times \omega_n$. Thus, ω is a *Mixed Nash Equilibrium* if, for any agent and distribution ω'_i , we have: $\mathbb{E}_{b \sim \omega}[u_i(b)] \geq \mathbb{E}_{b \sim (\omega'_i, \omega_{-i})}[u_i(b)]$.

3.3 The Prices of Stability and Anarchy

As our aim is to maximise the social welfare, we state that the allocation algorithm \mathcal{A} is a *c-approximation* algorithm if we have $SW(\mathcal{A}(v), v) \geq \frac{1}{c} SW_{opt}(v)$, where we denote $SW(\mathcal{A}(v), v)$ to represent the social welfare of the matching resulting from the allocation algorithm \mathcal{A} , and $SW_{opt}(v)$ for $\max_{M \in M(G)} SW(M, v)$ to represent the value of an optimal matching (and hence an optimal alignment) that maximises social welfare given the declaration vector v .

The trade-off between approximate (i.e. non-optimal) solutions and the optimal solution when identifying a matching is quantified as the *Price of Anarchy* [2]; i.e the ratio of the maximal possible social welfare and the social welfare emerging from an approximate solution. It is important to characterise this ratio as it provides a bound on how close an approximate algorithm can be to the optimal solution. The *Price of Anarchy* of the mechanism $\mathcal{M}(\mathcal{A}, \mathcal{P})$ in mixed (and pure, respectively) strategies can thus be defined as:

$$PoA_{mixed} = \sup_{v, \omega} \frac{SW_{opt}(v)}{\mathbb{E}_{b \sim \omega}[SW(\mathcal{A}(b), v)]}$$

$$PoA_{pure} = \sup_{v, b} \frac{SW_{opt}(v)}{SW(\mathcal{A}(b), v)}$$

where the supremum is over all valuations v and all mixed Nash Equilibria ω (likewise, all pure Nash Equilibria b) for v . Here, $\mathcal{A}(\omega)$ denotes a random matching with respect to ω .

⁶ The notion of a bid profile across a set of agents that omits the bid of agent i , represented as b_{-i} originates from the definition of the *Vickrey Clarke Groves* (VCG) mechanism [25], used extensively in mechanism design.

The *Price of Stability* is the ratio of the best stable matching with respect to the optimal matching. A bipartite graph may generate a number of sub-optimal but stable solutions; for example the classic Stable Marriage algorithm [14] typically generated matchings where the initial solution was optimal for one agent and yet pessimal for the other. The Price of Stability is important from a Mechanism Design perspective, as a mechanism (such as that discussion in Section 5) may compute the best stable solution and suggest it to the agents, who would implement this solution since it is stable. Thus, the price of stability captures this notion of optimisation subject to the stability constraint [2]. The price of stability for pure strategy games defined by mechanism $\mathcal{M}(\mathcal{A}, \mathcal{P})$ is the ratio between the best objective function value of one of its equilibria and that of the optimum:

$$PoS_{pure} = \inf_{v, b} \frac{SW_{opt}(v)}{SW(\mathcal{A}(b), v)}$$

where the infimum is over all type valuations v , and all pure Nash equilibria b .

4 Analysis of Alignment Selection with Payment

In this setting, we model the scenario as if both agents have to pay money to establish a matching (or ontological alignment), where the total cost is based on the bids declared for each correspondence. An agent may be incentivised to falsely lower the value of a correspondence, although this could result in it being rejected. Conversely, it may artificially inflate the value of the correspondence in the hope of it being selected; this however could result in a weaker, or inaccurate alignment. The aim here is to devise a mechanism that incentivises agents to be truthful when proposing correspondences, and to understand its properties.

The first observation is that this problem can be solved optimally using the *Vickrey Clarke Groves* (VCG) mechanism with Clarke payment [25]; which has the property that bidders can do no better than to bid their true valuations. In this analysis, we show that it is *not possible to have a faster, non-optimal, approximate and truthful mechanism for our problem*. This can be proved using the following lemma from classic mechanism design theory [25]:

Lemma 1. *An allocation rule of mechanism \mathcal{A} satisfies weak monotonicity if for all i and all v_{-i} , $\mathcal{A}(v_i, v_{-i}) = a \neq b = \mathcal{A}(v'_i, v_{-i})$ implies that $v_i(a) - v_i(b) \geq v'_i(a) - v'_i(b)$. If a mechanism $\mathcal{M}(\mathcal{A}, \mathcal{P})$ is incentive compatible, then \mathcal{A} satisfies weak monotonicity [25].*

The aim of Theorem 1 (below) is to determine if there is a mechanism that is not equivalent to VCG, yet is truthful, and to examine the quality of its solution. This theorem states that if any mechanism is not VCG, then: 1) it is not truthful; or 2) it is truthful but cannot achieve a solution whose approximation factor is smaller than 2.

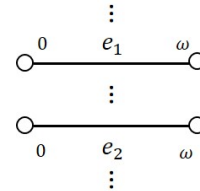


Fig. 4. Disjoint Edges

Theorem 1. *For the alignment problem with payment, any mechanism which does not adopt an optimal solution when agents declare their true valuations is either non-truthful, or if truthful, the non-optimal solution has an approximation ratio of at least 2.*

Proof. Let $\mathcal{M}(\mathcal{A}, \mathcal{P})$ be a mechanism, and recall that $\mathcal{A}(v)$ denotes the outcome generated by \mathcal{M} , when the input is a bid v (which may not be the true valuation).

Consider a bipartite graph of arbitrary size, where for two positive integers ℓ, k , let the bipartite graph $G = (U \cup V, E)$ have ℓ nodes on the left side of bipartite graph ($|U| = \ell$) and k nodes on the right side ($|V| = k$). We assume the existence of two special edges $e_1, e_2 \in E$ that are disjoint (i.e. $e_1 \cap e_2 = \emptyset$), such that their true valuations are $v_L(e_1) = v_L(e_2) = 0$ and $v_R(e_1) = v_R(e_2) = \omega$ (as illustrated in Figure 4). As the valuations of all other edges in G are zero for both agents (and thus do not appear in the Figure), the optimal solution should contain both edges e_1 and e_2 . As we only consider the problem from the perspective of the right agent in the discussion below, we omit the agent index when referring to valuations for simplicity.

Consider some mechanism $\mathcal{M}(\mathcal{A}, \mathcal{P})$ that generates a non-optimal solution which contains at most one of these edges. If neither e_1 and e_2 appear within the solution, the approximation ratio will be unbounded. Therefore, we assume that solution includes one of these two edges; w.l.o.g., assume that \mathcal{M} will accept $e_1 \in \mathcal{A}(v)$ when the right agent declares its true valuation v . If the right agent deviates from its valuation v to some other valuation v' , the mechanism has two options:

Case-1. The mechanism changes the current outcome to include both edges, such that original solution, $\mathcal{A}(v) \supseteq \{e_1\}$, is replaced with the solution, $\mathcal{A}(v') \supseteq \{e_1, e_2\}$. If we make the alternative valuation $v'(e_1) = v'(e_2) = 0$, this implies that $v'(\mathcal{A}(v')) \leq v'(\mathcal{A}(v))$. We also know that $v(\mathcal{A}(v)) < v(\mathcal{A}(v'))$. By adding the left and right hand sides of these two inequalities, we obtain:

$$\begin{aligned} v'(\mathcal{A}(v')) + v(\mathcal{A}(v)) &< v'(\mathcal{A}(v)) + v(\mathcal{A}(v')) \\ v(\mathcal{A}(v)) - v(\mathcal{A}(v')) &< v'(\mathcal{A}(v)) - v'(\mathcal{A}(v')) \end{aligned}$$

As this violates the weak monotonicity condition in Lemma 1, it follows that \mathcal{M} is not a truthful mechanism.

Case-2. The outcome is not changed, i.e., $\mathcal{A}(v') = \mathcal{A}(v)$ when the agent deviates its valuation to v' . The approximation ratio (i.e. the ratio of the approximate optimal solution to the optimal one) is at least $\frac{v(e_1) + v(e_2)}{v(e_1)}$. Since we consider the worst case, the ratio is at least 2 (i.e. the optimal solution is guaranteed to be within a factor of 2 of the returned solution).

If the outcome changes from $e_1 \in \mathcal{A}(v)$ to $e_2 \in \mathcal{A}(v')$ and $e_1 \notin \mathcal{A}(v')$, then this case is symmetric to Case 2, and thus will also lead to a ratio of at least 2. Furthermore, if the right agent has only one non-zero value edge, and the valuation on the remaining edges is 0, then the approximation ratio is unbounded, and all such cases also lead to the lower bound on the approximation ratio. \square

The motivation for the next theorem (Theorem 2) is that if a mechanism is not VCG but is truthful, then it must be *maximal-in-range* [8], defined below:

Definition 1. *A mechanism is called maximal in range (MIR) if there exists a fixed subset R of all allocations (the range of the mechanism), such that for every possible input v , the mechanism outputs the allocation that maximizes the social welfare in R with respect to v . [8]*

If a mechanism selects an edge such that the resulting solution (or *allocation*) is not one that is maximal with respect to the bids, then an agent will be incentivised to declare a lower (untruthful) valuation for an edge that they want in the solution, as this dishonest strategy will result in a higher utility than one that relies on being honest for the same solution.

Theorem 2. *For the alignment problem with payment, any deterministic mechanism which does not adopt an optimal solution when agents declare their true valuation is either non-truthful, or is a maximal-in-range mechanism.*

Proof. Consider a bipartite graph $G = (U \cup V, E)$ which contains ℓ nodes on the left ($|U| = \ell$), and a single node on the right ($|V| = 1$), and where there are ℓ edges, such that each node on the left is connected to the single node on the right. Thus, any solution will contain only a single edge. Furthermore, we assume that the optimal solution is $\{e_1\}$. If a deterministic mechanism \mathcal{A} *does not* adopt the optimal solution; then the solution generated by \mathcal{A} will be a single edge in $\{e_2, \dots, e_\ell\}$, where the optimal solution is e_2 (i.e. $v(e_1) > v(e_2)$). If the agents deviate from bidding their true value, the mechanism has three options:

Case-1. The solution adopted by mechanism \mathcal{A} does not change as a result of the changed bid. Thus, if \mathcal{A} is truthful then it is equivalent to a maximal-in-range mechanism, whose range is $R = \{e_2\}$.

Case-2. The solution adopted by mechanism \mathcal{A} changes to $\{e_1\}$ (the optimal solution for E) for some bid v' : $\mathcal{A}(v') = \{e_1\}$. Therefore, given Lemma 1, the mechanism \mathcal{A} cannot be truthful.

To show this, suppose w.l.o.g. that the mechanism adopts e_2 for the bid v : $\mathcal{A}(v) = \{e_2\}$, and that one agent deviates from its valuation v to v' such that $v'(e_1) < v'(e_2)$. Given that we have $v(e_1) > v(e_2)$; by adding the left and right hand sides of these two inequalities, we have:

$$v(e_2) - v(e_1) < v'(e_2) - v'(e_1)$$

As we have $\mathcal{A}(v) = \{e_2\}$ and $\mathcal{A}(v') = \{e_1\}$, this contradicts the monotonicity condition from lemma 1, which states that: $v(e_2) - v(e_1) \geq v'(e_2) - v'(e_1)$.

Case-3. The solution adopted by mechanism \mathcal{A} changes to a single edge from $\{e_3, \dots, e_\ell\}$. In such a case, by Lemma 1, the same argument for Case-2 also applies for this case where the mechanism is not truthful as it violates the monotonicity condition. \square

By combining the two theorems 1 and 2, we have the following theorem:

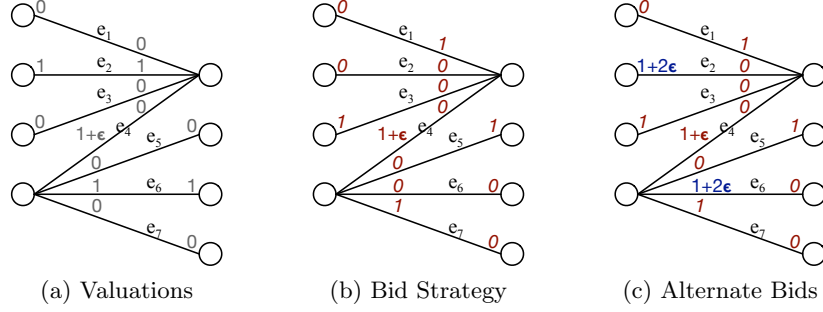


Fig. 5. Edge weights for the lower bound Price of Anarchy example

Theorem 3. *For the alignment problem with payment, the only truthful mechanisms are those that are maximal-in-range with an approximation ratio of at least 2.*

5 Nash equilibria implementation

Having analysed the Alignment Construction problem from a mechanism design perspective, we now explore the properties of a decentralised algorithm whereby two agents propose bids on candidate correspondences (not necessarily honestly) in order to determine a final alignment. In this section we explore a computationally efficient, yet sub-optimal setting using a *first price greedy matching* algorithm. This is a decentralised variant of the *NaiveDescending* algorithm given by Meilicke & Stuckenschmidt [22], and is presented in Algorithm 1. In this setting, the agents provide their declarations to the mechanism, which computes an outcome. The agents then measure their utility by subtracting their true valuation of this outcome by the payment. The payment scheme used models the notion that each agent would pay its own bid, i.e., $p_i = b_i(\mathcal{A}(b))$. The next two theorems provide a characterisation of the Price of Anarchy for a first-price greedy matching game. The proofs provide simple instances of the mechanism (from a game perspective) to give some intuition of pure Nash Equilibria.

Algorithm 1 Greedy algorithm

Require: Bipartite graph $G = (V \cup U, E)$, b_L, b_R are bids of the left & right agent.

Ensure: A matching M

Let $M = \emptyset$

if $E \neq \emptyset$ **then**

Find the edge $e \in E$ that maximises $b_L^e + b_R^e$

Let $M := M \cup \{e\}$

Remove from E edge e and edges incident to edge e

end if

M is the outcome

Theorem 4. *The price of anarchy (PoA) of the first price greedy matching game is at least 4.*

Proof. Consider a bipartite graph (Figure 5a), where the valuation for e_2 and e_6 are the same for both agents ($v_L^{e_2} = v_L^{e_6} = v_R^{e_2} = v_R^{e_6} = 1$); the valuation assigned to e_4 by the left agent is $v_L^{e_4} = 1 + \epsilon$ (where ϵ is a small positive number), whilst the right agent assigns this the value 0, and the remaining edges (i.e. e_1, e_3, e_5, e_7) have the valuation of 0 for both agents. Furthermore, assume a bid strategy profile (Figure 5b) for the left agent: $b_L^{e_4} = 1 + \epsilon, b_L^{e_3} = b_L^{e_7} = 1$; for the right agent it is: $b_R^{e_1} = b_R^{e_5} = 1$; and bids on the remaining edges being 0. Denote this strategy profile for both agents as b .

The greedy algorithm's solution given this profile b is $\{e_4\}$ (the sum of bids for e_4 is $1 + \epsilon$, whereas the sum for each of the other edges is either 0 or 1). The utility for the left agent given this solution is 0 (it would pay $1 + \epsilon$ because of its successful bid, but its valuation is $1 + \epsilon$), whereas the utility for the right agent is 0, as its bid and valuation for this edge are both 0.

The left agent could not unilaterally increase its utility; only one other solution $\{e_2, e_6\}$ has a positive utility, but to obtain this, new bids are necessary. If it chose two new bids (i.e. $\tilde{b}_L^{e_2}$ and $\tilde{b}_L^{e_6}$) on these edges such that $\tilde{b}_L^{e_2} > b_R^{e_1}$, which would result in the combined bids on e_2 being greater than that on e_1 (i.e. $v(e_2) > v(e_1)$) and $\tilde{b}_L^{e_6} > b_R^{e_5}$ (such that $v(e_6) > v(e_5)$), this solution would change to $\{e_2, e_6\}$ resulting in a negative utility for the left agent (Figure 5c). This is because its combined bid would be $2 + 4\epsilon$, whereas its payoff would be 2. The left agent will also not decrease its bid on e_4 , as the solution would be changed to another matching that is not $\{e_2, e_6\}$.

The right agent's behaviour is the same as the left, as this scenario is symmetric. It only has a positive valuation on e_2 and e_6 . By changing its bids for either edge, a new bid (e.g. $\tilde{b}_R^{e_2} > 1 + 2\epsilon$) would be required, thus again reducing the utility. Therefore in this case we have a Nash equilibrium, as neither agent can do better than adopt the current strategy.

The optimal solution is $\{e_2, e_6\}$, due to the joint valuation of $1 + 1 = 2$ for e_2 , and the same for e_6 , resulting in a total valuation of 4 for that solution. As stated above, the greedy algorithm instead finds the solution $\{e_4\}$, resulting in a total valuation of $1 + \epsilon$. Therefore, the Price of Anarchy is $\frac{4}{1+\epsilon}$, or at least 4. \square

Theorem 4 provides a lower bound on the price of anarchy for our mechanism. For the upper bound (Theorem 5), we first need the following two lemmas:

Lemma 2. *Suppose that the current bid profile (b_L, b_R) produces outcome M using a greedy mechanism. The necessary condition for (b_L, b_R) to be a Nash equilibrium is that $b_L^M \leq v_L^M$ and $b_R^M \leq v_R^M$.*

Proof. Assume that for outcome M , some agent's bid satisfies $b_i^M > v_i^M$ (i.e. the bid is greater than the valuation). The utility would then be $u_i(b_L, b_R) = v_i^M - b_i^M < 0$; i.e. it would be negative. Therefore, agent i would change its bid to a new one which increases its utility to a value that is at least 0. \square

Lemma 3. *Suppose that the current bid profile (b_L, b_R) produces an outcome M using a greedy mechanism, and $b_L^M \leq v_L^M$, $b_R^M \leq v_R^M$. There exists a bid for one agent, for example the left agent, \tilde{b}_L , that satisfies the condition $\tilde{b}_L^{M'} < 2(v_R^M + v_L^M) + \epsilon$. This would result in \tilde{b}_L changing the outcome to M' .*

Proof. Let $\{e_1, \dots, e_k\}$ be the set of edges in a matching M , indexed in decreasing order with respect to $b_L^e + b_R^e$. Denote e' as an edge in a different outcome M' . We assign each new bid $\tilde{b}_L^{e'}$ by the following procedure: $\forall j \in \{1, \dots, k\}$ (in this order), if the left side vertex of edge e_j has an adjacent edge e' in M' , then let the sum of the new bid (left) and the corresponding original bid (right) for M' take a slightly higher value than the corresponding edge bids for e_j in the outcome M ; i.e. $\tilde{b}_L^{e'} + b_R^{e'} > b_R^{e_j} + b_L^{e_j}$. Do the same for the right side vertex adjacent edge, i.e., for right side vertex adjacent edge $e' \in M'$ of e^j , let $\tilde{b}_L^{e'} + b_R^{e'}$ take a slightly higher value than $b_R^{e_j} + b_L^{e_j}$.

At any step of this procedure, if we need to reassign the bid $b_L^{e'}$ for some edge e' , then the bid of the larger value is retained (in fact, the declaration will remain unchanged as this procedure is conducted in decreasing order with respect to $b_L^e + b_R^e$). This distribution of bids is valid, as it can be done such that $\tilde{b}_L^{M'} > 2(b_R^M + b_L^M)$, which always results in a change of outcome to M' . It can also be easily argued that $\tilde{b}_L^{M'} < 2(v_R^M + v_L^M) + \epsilon$. \square

Theorem 5. *The price of anarchy (PoA) of a first price greedy matching game is at most 4.*

Proof. Let M be any matching whose total valuation is strictly smaller than a quarter of the optimum, i.e., $v_L^M + v_R^M < \frac{1}{4}Opt$. At least one of the following statements will hold on some other outcome M' given a different profile of valuations (either for the left or right agents respectively): $\exists M' v_L^{M'} \geq \frac{1}{2}Opt$ or $\exists M' v_R^{M'} \geq \frac{1}{2}Opt$. If M' is the optimal solution, then this will result in a contradiction. As they are symmetric, we assume the first statement is true. Assume $b = (b_L, b_R)$ is a fixed bid profile. If the outcome under b is M , then the agents will either have positive utilities; i.e. $b_L^M \leq v_L^M$ and $b_R^M \leq v_R^M$, or negative ones.

We want to show that the left agent would be incentivised to bid for the outcome M' . Let $\tilde{b}_L^{M'}$ be the bid that can achieve this change (i.e. from M to M'). Lemma 3 states that there exists some bid \tilde{b}_L that will achieve this change to outcome M' . Thus, we want to show that the utility of M' for the left agent is greater than for M ; i.e. $v_L^{M'} - \tilde{b}_L^{M'} > v_L^M - b_L^M$. By Lemma 3, since $\tilde{b}_L^{M'} < 2(v_R^M + b_L^M) + \epsilon$, we have:

$$v_L^{M'} - \tilde{b}_L^{M'} \geq v_L^{M'} - 2(v_R^M + b_L^M) - \epsilon$$

Since $v_L^{M'} \geq \frac{1}{2}Opt$ and $v_L^M + v_R^M < \frac{1}{4}Opt$, we can show that:

$$\begin{aligned} v_L^{M'} - 2(v_R^M + b_L^M) - \epsilon &\geq v_L^{M'} - (v_R^M + v_L^M) - v_R^M - b_L^M \\ v_L^{M'} - (v_R^M + v_L^M) - v_R^M - b_L^M &> v_L^M - b_L^M \end{aligned}$$

As ϵ can be arbitrarily small, it can be removed. The last inequality shows that the left agent can change its bid from b_L to \tilde{b}_L and get M' with a higher utility. This completes the argument as it shows that b cannot result in a Nash equilibrium. \square

Theorem 6. *The price of anarchy (PoA) of the first price greedy matching game is precisely 4.*

This theorem is the logical consequence of the theorems 4 and 5 that provide an upper and lower bound for the Price of Anarchy, so requires no further proof.

To conclude our analysis of the first-price greedy matching game, we investigate a lower bound for the Price of Stability through Theorem 7 (below).

Theorem 7. *The price of stability (PoS) of a first price greedy matching game is at least 2.*

Proof. Consider a bipartite graph (Figure 6), where the valuation assignment for both agents are: $v_L^{e_1} = v_L^{e_5} = 1$, $v_R^{e_2} = v_R^{e_4} = 1 + \epsilon$, $v_L^{e_3} = 1 + 3\epsilon$. The valuations on the remaining edges are 0 for both agents. The mechanism has three options:

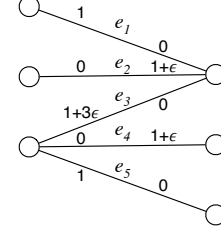


Fig. 6. Edge valuations for the PoS lower bound

Case-1. Suppose the outcome of the mechanism is $\{e_1, e_5\}$. The current bid cannot result in a Nash equilibrium, as the right agent would improve its utility by changing the current outcome to $\{e_2, e_4\}$, when $\tilde{b}_R^{e_2} > b_L^{e_1}$, $\tilde{b}_R^{e_4} > b_L^{e_5}$.

Case-2. Suppose the current outcome is $\{e_2, e_4\}$. It also does not admit any Nash equilibrium. If $\max\{b_R^{e_2}, b_R^{e_4}\} < v_L^{e_3}$, then the left agent could improve its utility by changing to e_3 , when $\tilde{b}_R^{e_3} > \max\{b_R^{e_2}, b_R^{e_4}\}$. If $\max\{b_R^{e_2}, b_R^{e_4}\} > v_L^{e_3}$, then let $b_R^{e_2}$ be a smaller bid, such that the left agent would then bid $\tilde{b}_L^{e_1} > b_R^{e_2}$ changing the outcome to $\{e_1, e_4\}$. This case is symmetric.

Case-3. Suppose the current outcome is $\{e_1, e_4\}$ (or $\{e_2, e_5\}$). The right agent would bid $\tilde{b}_R^{e_2} > b_L^{e_1}$ to improve its utility, and change the outcome to $\{e_2, e_4\}$.

To complete the proof, we provide a Nash equilibrium: $b_L^{e_1} = b_L^{e_5} = 1$, $b_R^{e_2} = b_R^{e_4} = 1 + \epsilon$, $b_L^{e_3} = 1 + 2\epsilon$. We can see in such a bid profile, the outcome would be e_3 , and it is easy to check that no agent can increase its utility. \square

It is usual in the literature to study the Price of Anarchy even if there might be instances without pure Nash equilibria [21]. Thus, Theorem 5 can be read as: *if there exists pure Nash equilibria, then their social welfare is at least 25% of the optimum.* We can also show that mixed Nash equilibria always exist, by transforming the problem into a new one in which each agent only has a finite number of strategies, where a strategy is for bids on edges. We define a small $\epsilon > 0$ as the minimum increment that any two bids can differ by. This leads to a finite number of strategies of any agent i as i will not bid more than $\sum_{e \in E} v_i(e)$. In particular, $b_i^e \in \{0, \epsilon, 2\epsilon, \dots, \sum_{e \in E} v_i(e)\}$.

Corollary 1. *The mixed Nash equilibrium exists for all instances of the discretised first price greedy matching game.*

This corollary is deduced directly from Nash’s theorem [25] which proves that if agents can use mixed strategies, then every game with a finite number of players in which each player can choose from finitely many pure strategies has at least one mixed Nash equilibrium.

Corollary 2. *The price of anarchy of the discretised first price greedy matching game for mixed strategy is 4.*

This proof can be found by extending that for Theorem 7.

6 Related Work

Approaches for resolving semantic heterogeneity have traditionally been centralised (i.e. with full access to the ontologies), resulting in the formation of a weighted bipartite graph representing the possible correspondences [22, 27], and a matching (alignment) found by either maximising social welfare or utilising a greedy search. However, these approaches were generally task agnostic, and thus varied in the way they utilised the weights. The lack of strategy or means to restrict what was revealed (due to knowledge encoded within an ontology being confidential or commercially sensitive) [12, 26] has resulted in an increased interest in decentralised, strategic approaches. Matchings have also been found through the use of *Argumentation*, based on private preferences over the correspondence properties (e.g., whether their construction was based on structural or linguistic similarities) [19], and public weights. More recently, dialogical approaches have been used to selectively exchange correspondences based on private weights for each agent [26]. Although polynomial approaches were used to determine the matching, the selection of revelations at each step was naive, and the resulting alignment failed represent the agents initial goals, whilst revealing the agents’ private weights.

Several studies have explored the problem of finding matchings from a mechanism design perspective, and have studied deterministic and randomise approximate mechanisms for bipartite matching problems where agents have one-sided preferences [17, 1]. Furthermore, there are a number of studies of truthful approximate mechanisms for combinatorial auctions, e.g., [5, 8, 20, 24], and various mechanisms [7, 11, 21] have studied Bayesian Nash Equilibrium settings. In [7], the problem of selling m items to n selfish bidders with combinatorial preferences, in m independent second-price auctions was studied. The authors showed that given submodular valuation functions, every Bayesian Nash equilibrium of the resulting game provided a 2-approximation to the optimal social welfare. The efficiency of Bayesian Nash equilibrium outcomes of simultaneous first- and second-price auctions was also studied [11], where bidders had complement-free (a.k.a. subadditive) valuations. They showed that the expected social welfare of any Bayesian Nash Equilibrium was at least $\frac{1}{2}$ of the optimal social welfare in the

case of first-price auctions, and at least $\frac{1}{4}$ in the case of second-price auctions. Lucier and Borodin [21] studied the general setting of combinatorial actions and proved that the Bayesian Price of Anarchy of the greedy algorithm is constant. A study of simultaneous second-price auctions [4] showed that the price of anarchy for pure Nash equilibrium was 2, and focused on Bayesian Nash equilibrium.

7 Conclusions

In this paper, we present, from a Mechanism Design perspective, the decentralised Ontology Alignment negotiation problem, whereby correspondences are selected for inclusion in an alignment (between two ontologies), and we provide a theoretical analysis of its properties. By demonstrating that different alignments can be generated depending on the selection process (e.g. by determining an optimal or sub-optimal solution), we characterise the problem analytically as a Mechanism Design problem, characterised as a Social Welfare maximising matching setting, where the valuation function is additive. We provide a complete picture of the complexity of this mechanism by showing that when coupled with a first-price payment scheme, it implements Nash equilibria which are very close (within a factor of 4) to the optimal matching. Furthermore, the *Price of Anarchy* of this mechanism is characterised completely and shown to be precisely 4 (this bound also holds for Mixed Nash equilibria), and when a pure Nash Equilibrium exists, we show that the *Price of Stability* is at least 2. Thus, decentralised agents can reach a Nash equilibrium, which produces a solution close to optimum within a factor of 4.

This analysis demonstrates that the type of alignment generated when selecting correspondences is sensitive to the algorithm used. However, by ensuring that the mechanism used is truth incentive, this ensures that agents will always do better by adopting strategies that accurately report the weights of their correspondences in decentralised settings.

References

1. Adamczyk, M., Sankowski, P., Zhang, Q.: Efficiency of truthful and symmetric mechanisms in one-sided matching. In: Algorithmic Game Theory, pp. 13–24 (2014)
2. Anshelevich, E., Das, S., Naamad, Y.: Anarchy, stability, and utopia: creating better matchings. *Autonomous Agents and Multi-Agent Systems* **26**(1), 120–140 (2013)
3. Atencia, M., Schorlemmer, W.M.: An interaction-based approach to semantic alignment. *Journal of Web Semantics* **12**, 131–147 (2012)
4. Bhawalkar, K., Roughgarden, T.: Welfare guarantees for combinatorial auctions with item bidding. In: Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms. pp. 700–709. SIAM (2011)
5. Briest, P., Krysta, P., Vöcking, B.: Approximation techniques for utilitarian mechanism design. In: STOC. pp. 39–48 (2005)
6. Chocron, P., Schorlemmer, M.: Vocabulary alignment in openly specified interactions. In: International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS). pp. 1064–1072 (2017)

7. Christodoulou, G., Kovács, A., Schapira, M.: Bayesian combinatorial auctions. In: *Int. Colloquium on Automata, Languages, and Programming*. pp. 820–832 (2008)
8. Dobzinski, S., Nisan, N., Schapira, M.: Approximation algorithms for combinatorial auctions with complement-free bidders. In: *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*. pp. 610–618. ACM (2005)
9. Euzenat, J.: Interaction-based ontology alignment repair with expansion and relaxation. In: *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017*. pp. 185–191 (2017)
10. Euzenat, J., Shvaiko, P.: *Ontology Matching, Second Edition*. Springer (2013)
11. Feldman, M., Fu, H., Gravin, N., Lucier, B.: Simultaneous auctions are (almost) efficient. In: *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*. pp. 201–210. ACM (2013)
12. Grau, B.C., Motik, B.: Reasoning over ontologies with hidden content: The import-by-query approach. *Journal of Artificial Intelligence Research* **45**, 197–255 (2012)
13. Gruber, T.R.: A translation approach to portable ontology specifications. *Knowledge Acquisition* **5**(2), 199–220 (1993)
14. Gusfield, D., Irving, R.W.: *The Stable Marriage Problem: Structure and Algorithms*. MIT Press, Cambridge, MA, USA (1989)
15. Jiménez-Ruiz, E., Payne, T.R., Solimando, A., Tamma, V.: Limiting consistency and conservativity violations through negotiation. In: *The 15th International Conference on Principles of Knowledge Representation and Reasoning (KR 2016)*. pp. 217–226 (2016)
16. Jimenez-Ruiz, E., Meilicke, C., Grau, B.C., Horrocks, I.: Evaluating mapping repair systems with large biomedical ontologies. In: *26th International Workshop on Description Logics (July 2013)*
17. Krysta, P., Manlove, D., Rastegari, B., Zhang, J.: Size versus truthfulness in the house allocation problem. In: *Proceedings of the fifteenth ACM conference on Economics and computation*. pp. 453–470. ACM (2014)
18. Kuhn, H.W.: The hungarian method for the assignment problem. *Naval research logistics quarterly* **2**(1-2), 83–97 (1955)
19. Laera, L., Blacoe, I., Tamma, V., Payne, T., Euzenat, J., Bench-Capon, T.: Argumentation over ontology correspondences in MAS. In: *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. pp. 1285–1292 (2007)
20. Lehmann, D.J., O’Callaghan, L., Shoham, Y.: Truth revelation in approximately efficient combinatorial auctions. *Journal of the ACM* **49**(5), 577–602 (2002)
21. Lucier, B., Borodin, A.: Price of anarchy for greedy auctions. In: *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*. pp. 537–553. Society for Industrial and Applied Mathematics (2010)
22. Meilicke, C., Stuckenschmidt, H.: Analyzing mapping extraction approaches. In: *Proc. 2nd Int Conf on Ontology Matching*. pp. 25–36 (2007)
23. Mitra, P., Lin, P., Pan, C.: Privacy-preserving ontology matching. In: *AAAI Workshop on Context and Ontologies*. vol. WS-05-01, pp. 88–91 (2005)
24. Mu’alem, A., Nisan, N.: Truthful approximation mechanisms for restricted combinatorial auctions. *Games and Economic Behavior* **64**(2), 612–631 (2008)
25. Nisan, N., Roughgarden, T., Tardos, E., Vazirani, V.V.: *Algorithmic game theory*, vol. 1. Cambridge University Press Cambridge (2007)
26. Payne, T.R., Tamma, V.: Negotiating over ontological correspondences with asymmetric and incomplete knowledge. In: *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. pp. 517–524 (2014)
27. Shvaiko, P., Euzenat, J.: Ontology matching: State of the art and future challenges. *IEEE Trans. Knowl. Data Eng.* **25**(1), 158–176 (2013)